C2 Advanced — LucaMac

Titolo nota — 07/09/2018

Metodo Probabilistico

Cosa è una variabile aleatoria?

"È una cosa che prende un po' di valori con un po' di probabilità"

Esempio: Se prendiamo un dado non truccato e chiamo D_6 la variabile aleatoria relativa
$\forall k \in \{1, \dots, 6\}$ vale $\mathbb{P}(D_6 = k) = 1/6$

Supponiamo che i nostri valori siano in $\mathbb{N}$. Sia X v.a.
Dunque $\sum_{n=0}^{\infty} \mathbb{P}(X=n) = \sum_{n \in \mathbb{N}} \mathbb{P}(X=n) = 1$

oss: possono anche essere tutti > 0. (Ad esempio $\mathbb{P}(X=n) = \frac{1}{2^{n+1}}$ oppure $\mathbb{P}(Y=n) = \frac{1}{(n+1)^2} \cdot \frac{6}{\pi^2}$).

oss2: Una condizione necessaria è $\lim_{n \to +\infty} \mathbb{P}(X=n) = 0$

oss3 (che doveva essere la prima) $\mathbb{P}(X=n) \in [0,1] \ \forall n \in \mathbb{N}$

oss4: Dato $A \subseteq \mathbb{N}$ vale $\mathbb{P}(X \in A) = \sum_{n \in A} \mathbb{P}(X=n)$

Esempio: consideriamo la X di oss1. E voglio $\mathbb{P}(X \in \{\text{numeri pari}\}) =$
$= \sum_{n \in \mathbb{N}} \mathbb{P}(X = 2n) = \sum_{n \in \mathbb{N}} \frac{1}{2 \cdot 4^n} = \frac{1}{2} \cdot \frac{1}{1 - \frac{1}{4}} = 2/3$ (< 1 😊)

Def: Data X variabile aleatoria definisco il suo valore atteso come
$$\mathbb{E}[X] = \sum_{n \in \mathbb{N}} n \, \mathbb{P}(X=n) \in [0, +\infty]$$
c'è $+\infty$

OSS: se $X \leq cost \Rightarrow \mathbb{E}[X] \leq cost \; (< +\infty)$

Esempio stupido: $\mathbb{E}[D_6] = \frac{1}{6} \sum_{n=1}^{6} n = \frac{7}{2} = 3.5$

OSS: se $X \geq Y$ v.a. $\Rightarrow \mathbb{E}[X] \geq \mathbb{E}[Y]$

Esempio concreto: Lancio due dadi $Y =$ primo valore $X =$ somma dei valori

$X \geq Y \Rightarrow \mathbb{E}[X] \geq \mathbb{E}[Y]$

(Per fare OSS serve la "vera" definizione di $X : \Omega \to \mathbb{N}$)

—o—o—

Indipendenza di 2 v.a.

Intuitivamente è $\forall n, k \in \mathbb{N} \quad \mathbb{P}(X = n \mid Y = k) = \mathbb{P}(X = n)$

Ovvero il valore di Y non cambia la distribuzione di X.

Formalmente è $\forall A, B \subseteq \mathbb{N} \quad \mathbb{P}(X \in A, Y \in B) = \mathbb{P}(X \in A)\mathbb{P}(Y \in B)$

OSS: $\mathbb{P}(X \in A \mid Y \in B) = \frac{\mathbb{P}(X \in A, Y \in B)}{\mathbb{P}(Y \in B)}$

Quindi X, Y indipendenti $\Leftrightarrow \forall A, B \subseteq \mathbb{N} \quad \mathbb{P}(X \in A \mid Y \in B) = \mathbb{P}(X \in A)$.

Lemma:

Se X ed Y sono indipendenti $\Rightarrow \mathbb{E}[X+Y] = \mathbb{E}[X] + \mathbb{E}[Y]$

Dim: Facile.

Lemma più figo: $\mathbb{E}[X+Y] = \mathbb{E}[X] + \mathbb{E}[Y]$

Cosa vuol dire $X+Y$?

Diamo la definizione "vera". (Con un errore)

X è una v.a. se è una funzione $X : \Omega \to \mathbb{N}$

$X(\omega) \in A \Leftrightarrow \omega \in X^{-1}(A)$

(Se $f : A \to B$ e $C \subseteq B$ $\quad f^{-1}(C) = \{a \in A : f(a) \in C\}$)

Allora $(X+Y)(\omega) := X(\omega) + Y(\omega)$.

“Dim Lemma + prop 2”

$$\mathbb{E}[X+Y] = \sum_{n=0}^{\infty} n\,\mathbb{P}(X+Y=n) = \sum_{n=0}^{\infty}\sum_{k=0}^{n} n\,\mathbb{P}(X=k,\, Y=n-k) =$$

$$= \sum_{n=0}^{\infty}\sum_{k=0}^{n} n\,\mathbb{P}(X=k)\cdot\mathbb{P}(Y=n-k\,|\,X=k) \underset{\text{tutto}\,\geq 0}{=} \sum_{k=0}^{\infty}\sum_{n=k}^{\infty} n\,\mathbb{P}(X=k)\cdot\mathbb{P}(Y=n-k\,|\,X=k) =$$

$$\overset{m=n-k}{=} \sum_{k=0}^{\infty}\sum_{m=0}^{\infty} (m+k)\,\mathbb{P}(X=k)\,\mathbb{P}(Y=m\,|\,X=k) = \sum_{m=0}^{\infty}\sum_{k=0}^{\infty} (m+k)\,\mathbb{P}(X=k)\cdot\mathbb{P}(Y=m\,|\,X=k) =$$

$$= \sum_{m=0}^{\infty}\left(m \underbrace{\sum_{k=0}^{\infty} \mathbb{P}(X=k)\cdot\mathbb{P}(Y=m\,|\,X=k)}_{\mathbb{P}(Y=m)}\right) + \sum_{k=0}^{\infty}\left(k\,\mathbb{P}(X=k)\underbrace{\sum_{m=0}^{\infty}\mathbb{P}(Y=m\,|\,X=k)}_{1}\right) =$$

$$= \sum_{m=0}^{\infty} m\,\mathbb{P}(Y=m) + \sum_{k=0}^{\infty} k\,\mathbb{P}(X=k) = \mathbb{E}[X] + \mathbb{E}[Y]$$

Tutto quello che abbiamo fatto funziona in modo analogo se al posto di $\mathbb{N}$ ci metto un qualunque $A \subseteq \mathbb{R}_{\geq 0}$ numerabile.

Disuguaglianza di Markov

Se $X \geq 0$ v.a. e $c > 0$. Allora

$$\mathbb{P}(X \geq c) \leq \frac{\mathbb{E}[X]}{c}$$

Dim: $I_{\{X\geq c\}} = \begin{cases} 1 & \text{se } X \geq c \\ 0 & \text{se } X < c \end{cases}$

Vale $c\,I_{\{X\geq c\}} \leq X$ perché $\overset{x\geq c}{\nearrow}\; X \geq c$, $\underset{x<c}{\searrow}\; X \geq 0$

Allora $\mathbb{E}[X] \geq \mathbb{E}[c\,I_{\{X\geq c\}}] = c\cdot\mathbb{P}(X\geq c) + 0\cdot\mathbb{P}(X<c) = c\,\mathbb{P}(X\geq c)$

$$\Rightarrow \mathbb{P}(X \geq c) \leq \frac{\mathbb{E}[X]}{c}$$

Def: Si dice varianza di X $\quad Var(X) := \mathbb{E}[|X - \mathbb{E}[X]|^2] = \mathbb{E}[X^2] - \mathbb{E}[X]^2$

$LHS = \mathbb{E}[X^2] - \cancel{2}\,\mathbb{E}[X]^2 + \cancel{\mathbb{E}[X]^2}$

Disuguaglianza di Chebishev

X v.a. (ma consideriamo ancora $X \geq 0$) e $c > 0$

$$\mathbb{P}(|X-\mathbb{E}[X]| \geq c) \leq \frac{Var(X)}{c^2}$$

OSS: $|X-\mathbb{E}[X]| \geq 0$

MARKOV su $|X-\mathbb{E}[X]|^2 \geq 0$ e $c^2 > 0$

$$\mathbb{P}(|X-\mathbb{E}[X]| \geq c) = \mathbb{P}(|X-\mathbb{E}[X]|^2 \geq c^2) \leq \frac{\mathbb{E}[|X-\mathbb{E}[X]|^2]}{c^2} = \frac{Var(X)}{c^2}$$

—o—o—

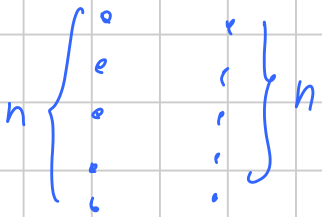
Calcolare il valore atteso del # di pt fissi in una permutazione di $\{1,\dots,n\}$

$$X_i(\sigma) = \begin{cases} 1 & \text{se } i \text{ punto fisso di } \sigma \\ 0 & \text{altrimenti} \end{cases}$$

$X_i : S_n \to \{0,1\}$

$$\mathbb{E}[X_1 + \dots + X_n] = \sum_{i=1}^{n} \mathbb{E}[X_i] = \sum_{i=1}^{n} \frac{(n-1)!}{n!} = \sum_{i=1}^{n} \frac{1}{n} = n \cdot \frac{1}{n} = 1$$

Esempi:



Grafo bipartito e ci sono almeno n^2-n+1 lati. Tesi: $\exists$ perfect matching.

Scelgo n coppie disgiunte in modo random

$$X_i = \begin{cases} 1 & \text{se collegati} \\ 0 & \text{altrimenti} \end{cases}$$

Se avessi $\sum_{i=1}^{n} X_i = n$ una volta avrei vinto.

$$\mathbb{E}\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} \mathbb{E}[X_i] \geq \sum_{i=1}^{n} \frac{n^2-n+1}{n^2} = \frac{n^2-n+1}{n} = n-1+\frac{1}{n}$$

$\sum_{i=1}^{n} X_i$ è intero $\Rightarrow$ almeno una volta è n

Es 2 n numeri reali $\sum_{i=1}^{n} a_i = 0$ e non sono tutti 0.

Tesi: $\exists$ permutazione di questi $b_1, \dots, b_n$ t.c. $b_1b_2 + \dots + b_nb_1 < 0$

$$\mathbb{E}\left[\sum_{i=1}^{n}{}' b_ib_{i+1}\right] = \sum_{i=1}^{n}{}' \mathbb{E}[b_ib_{i+1}] = n\,\mathbb{E}[b_1b_2] = n \cdot \frac{\sum_{i\neq j}' a_ia_j}{\frac{n(n-1)}{2}} =$$

$$= \frac{2}{n-1}\sum_{i<j}{}' a_ia_j = \frac{1}{n-1}\left(\left(\sum_i{}' a_i\right)^2 - \sum_i{}' a_i^2\right) = -\frac{\sum a_i^2}{n-1} < 0$$

Da qui è easy.

Nuova tecnica: IMC 2017

In una città ci sono n persone. Ogni persona conosce esattamente 1000 persone.
Dimostrare che $\exists$ S insieme di persone t.c. almeno $\frac{n}{2017}$ persone in S conoscano esattamente 2 persone in S.

Ogni persona scelgo (indipendentemente) con probabilità p di metterla in S

$$X_i = \begin{cases} 1 & \text{se } i \in S \text{ ed esattamente 2 amici sono in } S \\ 0 & \text{altrimenti} \end{cases}$$

Voglio che almeno una volta $\sum X_i \geq \frac{n}{2017}$, quindi mi basta

$$\mathbb{E}\left[\sum{}' X_i\right] \overset{?}{\geq} \frac{n}{2017}$$

$$\mathbb{E}\left[\sum{}' X_i\right] = \sum{}' \mathbb{E}[X_i] = \sum{}' p \cdot \binom{1000}{2} p^2 (1-p)^{998} = n\binom{1000}{2} \cdot p^3 \cdot (1-p)^{998}$$

$$\binom{1000}{2} p^3 (1-p)^{998} \overset{?}{\geq} \frac{1}{2017} \qquad p = \frac{3}{1001}$$

$$\binom{1000}{2} \cdot \left(\frac{3}{1001}\right)^3 \cdot \left(1 - \frac{3}{1001}\right)^{998} \geq \binom{1000}{2} \cdot \left(\frac{3}{1001}\right)^3 \cdot \underbrace{\left(1 - \frac{3}{1001}\right)^{1001}}_{\geq e^{-3}}$$

$$\frac{1000 \cdot 999}{2} \cdot \frac{1}{(1001)^3} \overset{?}{\geq} \frac{1}{2017}$$

$n \geq 1000$ $\quad (2n+17)\, n\, (n-1) \overset{?}{\geq} 2(n+1)^3$ fate il conto

$9n^2 \overset{?}{\geq} 23n + 2$ che è vero!

$\Rightarrow \mathbb{E}\left[\sum X_i\right] \geq \dfrac{e^{-3} \cdot 27}{2017} \geq \dfrac{1}{2017}$ Tesi

(4) G grafo con n vertici e grado medio $d \geq 1$.

$\Rightarrow \exists$ anticricca con almeno $\frac{n}{2d}$ vertici.

Prendo ogni vertice con probabilità p. $\mathbb{E}[\#\text{vertici}] = np$

Considero un lato. Qual è la prob che entrambi i vertici siano presi? p^2

$\mathbb{E}[\#\text{lati t.c. ci sono entrambi i vertici}] = p^2 \cdot \#\text{lati} = \frac{1}{2} n d p^2$

Tolgo un vertice per ogni lato con due vertici presi

$\mathbb{E}[\#\text{vertici tolti}] \leq \frac{1}{2} n d p^2$

$$p = \frac{1}{d} \leq 1$$

$\mathbb{E}[\#\text{vertici restanti}] \geq np - \frac{1}{2} n d p^2 \overset{\uparrow}{=} \frac{n}{d} - \frac{1}{2}\frac{n}{d} = \frac{n}{2d}$

Ora {vertici restanti} formano un'anticricca

$\Rightarrow \exists$ anticricca con almeno $\frac{n}{2d}$ vertici.

(5) Stessa cosa, senza l'ipotesi $d \geq 1$, e $\sum \frac{1}{d_i + 1}$

Consideriamo una permutazione random σ (tutte prob $= 1/n!$)

Sia A_i l'evento $\sigma(i) < \sigma(j)\ \forall j$ vicino ad i $\Rightarrow \mathbb{P}(A_i) = \dfrac{1}{d_i + 1}$

$\quad$ " $\{\sigma :$ vale $\sigma(i) < \sigma(j)\ \forall j$ vicino ad $i\}$

$$X(\sigma) = \sum_{i \geq 1}^{n} I_{\{\sigma \in A_i\}}$$

Se $\exists\, \sigma$ t.c. $X(\sigma) \geq m \Rightarrow \exists$ anticricca grande m

Basta $\mathbb{E}[X] \overset{!}{\geq} \sum' \frac{1}{d_i+1}$

$$\mathbb{E}[X] = \sum' \mathbb{E}[I_{\sigma \in \{A_i\}}] = \sum' \mathbb{P}(\sigma \in A_i) = \sum' \mathbb{P}(A_i) = \sum' \frac{1}{d_i+1}$$

OSS: Per AM-HM $\sum' \frac{1}{d_i+1} \geq \frac{n^2}{(\sum d_i)+n} = \frac{n}{d+1} \underset{d \geq 1}{\geq} \frac{n}{2d}$

6) USAMO 2012.6

$n \geq 2$ $x_1, \dots, x_n \in \mathbb{R}$ con $x_1 + \dots + x_n = 0$ $x_1^2 + \dots + x_n^2 = 1$

$\forall A \subseteq \{1, \dots, n\}$ definiamo $S_A = \sum'_{i \in A} x_i$

$\forall \lambda > 0$ $\exists$ al più $\frac{2^{n-3}}{\lambda^2}$ sottoinsiemi A con $S_A \geq \lambda$.

7) RUSSIA QUALCOSA

Ad ogni ragazzo piace almeno una ragazza.

Dimostrare che $\exists\, S$ con almeno la metà delle persone nel mondo t.c. ad ogni ragazzo in S piacciono un numero dispari di ragazze in S

8) IMO 1998. qualcosa

In una competizione ci sono a partecipanti e b giudici con $b \equiv_2 1$. Ogni giudice dice V o X $\forall$ partecipante. Sappiamo che $\forall$ coppia di giudici # partecipanti con lo stesso "voto" è $\leq k$. Dimostrare che $k \geq \frac{a(b-1)}{2b}$.

9) (quasi) IMO 2014.6

n rette in posizione generale. Dimostra che per n abb grande si possono colorare almeno $\sqrt{n}$ rette di viola in modo tale che nessuna regione finita abbia tutti i bordi colorati

(Nota: punti parziali per $c\sqrt{n}$ con $c < 1$)

Fatelo $\forall c < \frac{2}{3}$

6) $X : \mathcal{P}(\{1,\dots,n\}) \to \{\text{un po' di valori}\}$

$X(A) = S_A$ (scriverò S_A)

$\mathbb{E}[S_A] = 0$ perché

$S_A + S_{A^c} = 0$

$\# S_A$ t.c. $\geq \lambda$ è $\leq \dfrac{2^{n-3}}{\lambda^2}$

ovvero $\mathbb{P}(S_A \geq \lambda) \overset{?}{\leq} \dfrac{1}{8\lambda^2}$

$\mathbb{P}(S_A \geq \lambda) = \frac{1}{2}\mathbb{P}(|S_A| \geq \lambda)$ perché $\lambda > 0$ e $S_{A^c} = -S_A$.

Dunque basta $\mathbb{P}(|S_A| \geq \lambda) \leq \dfrac{1}{4\lambda^2}$

Ora vorrei usare Chebishev:

per Chebishev

$$\mathbb{P}(|S_A| \geq \lambda) \leq \frac{Var(S_A)}{\lambda^2}$$

$$Var(S_A) = \mathbb{E}[S_A^2] \cancel{\left(-\mathbb{E}[S_A]^2\right)}$$

Ora $\mathbb{E}[S_A^2] = \dfrac{1}{2^n}\displaystyle\sum_{A \subseteq \{1,\dots,n\}}' \left(\left(\sum_{i \in A}' x_i\right)^2\right) = \dfrac{1}{2^n}\sum_{A \subseteq \{1,\dots,n\}} \left(\sum_{i \in A}' x_i^2 + 2\sum_{\substack{i,j \in A \\ i<j}}' x_i x_j\right) =$

$= \dfrac{1}{2^n}\left[\displaystyle\sum_{i=1}^{n}' 2^{n-1} x_i^2 + \sum_{i<j}' 2^{n-1} x_i x_j\right] = \dfrac{1}{2}\left(\sum x_i^2 + \dfrac{1}{2}\left(\left(\sum' x_i\right)^2 - \sum' x_i^2\right)\right) =$

$= \dfrac{1}{4}$

$\mathbb{P}(|S_A| \geq \lambda) \leq \dfrac{1}{4\lambda^2}$ che è quello che volevamo.

7) Scegliamo con probabilità $1/2$ (ed in modo indipendente) ogni rofetto, e prendiamo tutti i rofetti a cui ne piacciono un # dispari.

Sia $G = \#$ girls $\quad B = \#$ boys

$$\mathbb{E}[\#\text{ ragazze prese}] = \frac{G}{2}$$

$$\mathbb{E}[\#\text{ ragazzi presi}] = \sum_{\text{ragazzi}} \mathbb{P}(\text{preso}) = \sum_{\text{ragazzi}} \frac{1}{2} = \frac{B}{2}$$

$$\mathbb{E}[\#\text{ persone prese}] = \frac{B+G}{2}$$

Dunque $\exists$ configurazione in cui ne prendo almeno $\frac{B+G}{2}$.

8) C contestant random (in maniera equiprobabile)

Ora v ed x $\# v$ e $\# x$.

Sappiamo $v + x = b$

Ora quanti i sono i giudici che concordano su C?

$$\binom{v}{2} + \binom{x}{2} \overset{?}{\geq} \frac{(b-1)^2}{4}$$

$$\iff v^2 + x^2 - (v+x) \overset{?}{\geq} \frac{(b-1)^2}{2} = \frac{b^2+1}{2} - b$$

$$\iff v^2 + x^2 \overset{?}{\geq} \frac{b^2+1}{2}$$

Ma $v^2 + x^2 \geq \frac{(b-1)^2}{4} + \frac{(b+1)^2}{4} = \frac{b^2+1}{2}$ (oppure Q.M.-A.M. + $\nexists$ =)

$$\frac{1}{a} \cdot |I| = \mathbb{E}\left[\binom{v}{2} + \binom{x}{2}\right] \geq \frac{(b-1)^2}{4}$$

dove $I = \{(\alpha, \beta, \gamma)$ con α, β giudici γ partecipante e α, β hanno dato lo stesso voto a $\gamma\}$

$$|I| \leq K\binom{b}{2}$$

$$\Rightarrow K \geq \frac{a(b-1)^2}{4} \cdot \frac{2}{b(b-1)} = a\frac{(b-1)}{2b}$$

Tesi!

9) IMO 2014.6

Coloro ogni retta con prob p

$\mathbb{E}[\#\text{rette colorate}] = np$

Quante sono le aree finite (al più)? $\leq \frac{n^2}{2}$

Allora $\mathbb{E}[\#\text{aree cattive}] \leq \frac{n^2}{2} p^3$

Al solito tolgo 1 retta $\forall$ aree cattive

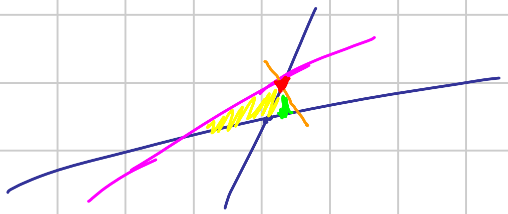
E viene

$$\mathbb{E}[\#\text{rette colorate restanti}] \geq np - \frac{n^2 p^3}{2}$$

$n = \frac{3}{2} n^2 p^2 \quad p = \sqrt{\frac{2}{3n}}$

$$\mathbb{E}[\quad - \quad - \quad - \quad] \geq \sqrt{n}\left[\sqrt{\frac{2}{3}} - \frac{1}{2} \cdot \frac{2}{3} \cdot \sqrt{\frac{2}{3}}\right] = \sqrt{n} \cdot \left(\frac{2}{3}\right)^{3/2}$$

Facciamo una cosa un po' più raffinata



Voglio contare $\max\{\# \triangle\}$.

Ogni vertice è al più in 2 triangoli (dim audio e figura)

$\Rightarrow \# \triangle \leq \#\text{vertici} \cdot \frac{2}{3} \leq \frac{n^2}{3}$

Quindi faccio come prima, solamente se non $\triangle$ ci metto p^4

Come prima, stimando gli "almeno quadrilateri" come al più $n^2/2$, otteniamo

$$np - \frac{n^2}{3}p^3 - \frac{n^2}{2}p^4 \qquad p = \frac{k}{\sqrt{n}}$$

$$\underbrace{\sqrt{n}\left(k - \frac{k^3}{3}\right)}_{k=1} - \frac{k^4}{2}$$

$$\frac{2}{3}\sqrt{n} - \frac{1}{2}.$$

Allora, visto che $\forall c < \frac{2}{3}$ per n abbastanza grande si ha che

$$\frac{2}{3}\sqrt{n} - \frac{1}{2} \geq c\sqrt{n}$$

$\Rightarrow$ Ho la tesi $\forall c < \frac{2}{3}$.

C'è un metodo simile, ma più complicato, attraverso il quale si arriva alla stima $c\sqrt{n\log n}$ per un qualche $c > 0$.